



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

A

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/513,058	02/24/2000	Anita Wai-Ling Huang	AM999074	5519
23334	7590	07/26/2005	EXAMINER	
FLEIT, KAIN, GIBBONS, GUTMAN, BONGINI & BIANCO P.L. ONE BOCA COMMERCE CENTER 551 NORTHWEST 77TH STREET, SUITE 111 BOCA RATON, FL 33487			BASEHOAR, ADAM L	
ART UNIT		PAPER NUMBER		2178
DATE MAILED: 07/26/2005				

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/513,058	HUANG ET AL.	
	Examiner	Art Unit	
	Adam L. Basehoar	2178	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) Responsive to communication(s) filed on 02 May 2005.

2a) This action is **FINAL**. 2b) This action is non-final.

3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) Claim(s) 1,2,5-12,14,17,18,20,21,25 and 26 is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5) Claim(s) _____ is/are allowed.

6) Claim(s) 1,2,5-12,14,17,18,20,21,25 and 26 is/are rejected.

7) Claim(s) _____ is/are objected to.

8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) The specification is objected to by the Examiner.

10) The drawing(s) filed on _____ is/are: a) accepted or b) objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a) All b) Some * c) None of:

1. Certified copies of the priority documents have been received.

2. Certified copies of the priority documents have been received in Application No. _____.

3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) Notice of References Cited (PTO-892)

2) Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date _____.

4) Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____.

5) Notice of Informal Patent Application (PTO-152)

6) Other: _____.

DETAILED ACTION

1. This action is responsive to communications: The Amendment filed 05/02/05.
2. Claims 3-4, 13, 15-16, 19, and 22-24 have been cancelled by Amendment.
3. The rejection of claims 1-2, 5-9, 12, 14, 17-18, 20-21 remain under 35 U.S.C. 103(a) as being unpatentable over Brown et al (US: 5,913,208, 06/15/99).
4. The rejection of claims 10-11 and 25-26 remain under 35 U.S.C. 103(a) as being unpatentable over Brown et al (US: 5,913,208, 06/15/99) in view of Microsoft Press Computer Dictionary, Microsoft Press, 1997, pp. 309.
5. Claims 1, 2, 5-12, 14, 17, 18, 20, 21, 25, and 26 are pending in the case. Claims 1, 10, 12, 17, and 25 are independent claims.

Claim Rejections - 35 USC § 103

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

7. Claims 1-2, 5-9, 12, 14, 17-18, 20-21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Brown et al (US: 5,913,208, 06/15/99).

-In regard to substantially similar independent claims 1, 12, and 17, Brown et al teach a method, system, and computer program product comprising:
receiving a first and a second document (column 4, lines 59-66)(Fig. 3B);

a metadata parser generating a first and second metadata summary for each of the received documents (column 7, lines 40-67)(Fig. 3B), wherein metadata summaries included a plurality of sub-trees(i.e. Intrinsic and Non-Intrinsic Attributes)(column 7, lines 48-67)(Fig. 3B: 370 & 380) wherein each sub-tree could include a plurality of nodes (i.e. Intrinsic (relevance score, title, size) and Non-Intrinsic (location which includes filename (Fig. 2b));

storing the metadata summaries in a repository (Fig. 1 & Fig. 4: 410);

a summary consolidator for comparing the metadata summaries on a structural level by comparing the corresponding sub-tees (Fig. 4: 420, 430 & 440, 450);

identifying the first and second documents as distinct if the structure of the sub-trees are not equivalent (Fig. 4: 455);

if the structures are equivalent performing the steps of:

comparing the first and second metadata summaries on a textual level (i.e. by comparing relevance score (Fig. 3B: 375) which was calculated by information retrieval algorithm as a function of the query (Fig. 3A) and the contents of the document (column 7, lines 55-67)) by comparing the textual content from the first and second document contained in the metadata summaries (Fig. 4: 420); and

identifying the first and second documents as distinct if the textual content within the sub-trees are not equivalent (Fig. 4: 455).

Brown et al do not specifically teach comparing the metadata summaries on a structural level before comparing the textual content (e.g. “relevance score” “title”). It would have been obvious to one of ordinary skill in the art at the time of the invention for Brown et al to have compared the sub-tree structures of the metadata summaries of the documents to determine

document distinctness before comparing the textual content of the documents to determine distinctness, because by checking the structure of sub-trees first, it could more quickly be established with the minimum amount of comparisons that the compared documents were distinct from one another (e.g. Referring to Fig. 3B it is shown that there are fewer sub-tree elements than there are textual values within the elements. Thus it was obvious if one document maintained a “Title” structure and a compared document did not then those two documents would be distinct and said distinction would be obtained more quickly than parsing and comparing the one document title of “HEAVY RAINS FLOOD FARMS IN MID-WEST” with no title of another document)

-As per dependent claims 2, 14, and 18, Brown et al further disclose comparing the first and second metadata summaries on an attribute level (author, creation date, length, size, location, abstract, etc (columns 1 and 6, lines 64-67 and 51-54)) by comparing attribute values within the sub-trees of the first and second metadata summaries and identifying the first and second documents as distinct if the attribute values within the sub-trees are not equivalent. (Fig. 4: 455). Similar to the independent claims discussed above, Brown et al do not specifically teach wherein the attribute comparison was made after the structural comparison and before the textual comparison. Similarly, it would have been obvious to one of ordinary skill in the art at the time of the invention for Brown et al to have compared the attributes of the metadata summaries after the structural comparison and before the textual comparison, because an attribute comparison provided the second quickest determination of distinctness between the documents (i.e. checking the attribute values stored in the metadata summaries required more processing than checking a

few nodes in the metadata summaries and required less than checking the textual content of the document metadata summaries).

-As per dependent claims 5 and 20, Brown et al further disclose noting the current pair of hit-list documents as duplicates if the text content (column 7, lines 55-67) (Fig. 3B: 375) are equal (Fig. 4: 432) (i.e. the structure and non-text attributes have been compared but alone can not determine duplicates)(Fig. 4: 460).

-As per dependent claims 6 and 21, Brown et al further disclose one embodiment of *deleting duplicates* from the hit list group (column 8, lines 20-22: Fig. 8).

-As per dependent claims 7-9, Brown et al further disclose that in the hit-list (metadata table) when an entry is identified as having one or more duplicates.....is cross referenced in the duplicate identifier field of the other duplicate (column 8, lines 11-16)(Fig. 3b: 390), wherein for two metadata entries the *entry number and duplicate identifier field constitute a 2x2 matrix* (Fig. 3b). Brown et al also teach that the process of identifying duplicates as distinct is achieved by comparing the metadata structures, attributes, and text content of the attributes as seen in independent claim 1. Brown et al do not teach storing a zero binary value in the first row and second column position of the matrix to designate that the first and second documents are distinct. It would have been obvious to one of ordinary skill in the art, to have used a binary indicator (0 or 1) in Brown et al duplicate identifier position concerning only two documents, because while Brown et al is set up for a plurality of documents which could lead to many

multiple dependencies, if a restriction of only comparing two documents was applied to Brown et al, there would be no need for more than one indicator (0 or 1) to clearly state the relationship between the first document and second document.

8. Claims 10-11 and 25-26 are rejected under 35 U.S.C. 103(a) as being unpatentable over Brown et al (US: 5,913,208, 06/15/99) in view of Microsoft Press Computer Dictionary, Microsoft Press, 1997, pp. 309.

-As per independent claims 10 and 25, Brown et al teach the method of claim 1 as stated above, and further disclose receiving not just two, but *two or more entries* (column 3, lines 54-55) of documents and generating metadata hit-list summaries for each received document (Fig. 3b). Brown et al do not teach that when receiving a plurality of documents to group the metadata hit-list summaries according to the mime-type designation of each document. The Microsoft Press Computer Dictionary teaches us that MIME-types describe the contents of a document which can be used to interpret the content of a file over the internet (pp. 309). It would have been obvious to one of ordinary skill in the art, to have used Brown et al method for identifying duplicate documents from search results without comparing document content and grouping the hit-list summaries by MIME-type, because documents of different MIME-types wouldn't have to be compared as their intrinsic attributes would be wholly different, i.e. a regular text/plain MIME-type couldn't be a duplicate of a text/html MIME-type, and thus this would increase efficiency by significantly reducing the number of hit-list summary comparisons. Note the above rejections of independent claims 1, 12, and 17 with regards of comparing the metadata summaries on a structural level to determine document equivalence.

-As per dependent claims 11 and 26, Brown et al teach generating metadata summaries for each received document (Fig. 3B). Brown et al and the Microsoft Press Computer Dictionary do not teach grouping more subsets of hit-list summaries by MIME-type. As it was taught above in the rejection of independent claims 10 and 25, it would have been obvious to have grouped as many MIME-type subsets as were required by the plurality of documents for the purpose of increased efficiency by significantly reducing the number of hit-list summary comparisons across MIME-types.

Response to Arguments

9. Applicant's arguments filed 05/02/05 have been fully considered but they are not persuasive.

-In regard to independent claims 1, 12, and 17, Applicant generally argues that the Brown reference does not all features of the independent claims (Remarks: Page 13, 2nd Paragraph). The Examiner respectfully disagrees with the Applicant. The Examiner believes the Brown reference teaches generating metadata summaries for a plurality of documents (Fig. 3B). Within these metadata document summaries, Brown teach include a plurality of sub-trees (i.e. one sub-tree including intrinsic attributes and another sub-tree including non-intrinsic attributes) wherein each of the sub-trees include a plurality of nodes (i.e. Intrinsic (relevance score, title, size) and Non-Intrinsic (location which includes filename (Fig. 2b)). The Examiner agrees that Brown does not specifically teach first comparing the structures of the document metadata summaries to determine distinctness. However the Examiner as shown above believes that comparison would

have been well known in the art at the time of the invention for obvious reasons. Brown also teaches comparing the metadata summaries on a textual level via the relevancy score which takes into account the entire content of the document which would include textual content of the document (i.e. at least the title information provided in the metadata summaries). Similarly to independent claims 1, 12, and 17, the Examiner believes the Brown reference in view of the well known functionality of mime-types as shown in the computer dictionary would have been an obvious combination to reduce the number of the metadata summary comparisons.

The Examiner believes to overcome the current prior art certain features of the Applicant invention must be further refined. Possibilities include a more detailed recitation of the textual content or of the plurality of sub-trees including a plurality of nodes. In light of further amendments to the claims the Examiner points out the newly cited reference of Aoyama et al (US-5,965,726) which teaches both structural and textual comparison of documents to determine distinctness between the documents.

Conclusion

10. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period

Art Unit: 2178

will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

11. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

US-6,502,112	12/31/02	Baisley
US-5,965,726	09/21/99	Aoyama et al.

12. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Adam L. Basehoar whose telephone number is (571)-272-4121. The examiner can normally be reached on M-F: 7:00am - 4:00pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Steve Hong can be reached on (571) 272-4124. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Application/Control Number: 09/513,058

Art Unit: 2178

Page 10


STEPHEN HONG
SUPERVISORY PATENT EXAMINER

ALB